



Universidade de Lisboa
Faculdade de Motricidade Humana



SF-FAB: Fiabilidade teste-reteste e capacidade preditiva de quedas em idosos portugueses a viver na comunidade.

Dissertação elaborada com vista à obtenção do Grau de Mestre em Exercício e Saúde

Orientador: Doutora Vera Moniz Pereira da Silva

Júri:

Presidente:

Doutora Maria Filomena Araújo Costa Cruz Carnide, Professora Auxiliar
da Faculdade de Motricidade Humana da Universidade de Lisboa

Vogais:

Doutora Vera Moniz Pereira da Silva, Professora Auxiliar da Faculdade
de Motricidade Humana da Universidade de Lisboa

Doutor Eduardo José Brazete Carvalho Cruz, Professor Coordenador da
Escola Superior de Saúde do Instituto Politécnico de Setúbal

Guilherme Strelow Hilger

2019

Agradecimentos

Em primeiro lugar, um enorme agradecimento à orientadora da minha dissertação, Professora Doutora Vera Moniz Pereira, por me ajudar de todas as formas possíveis a dar continuidade em um projeto que já estava em andamento. Acrescento um agradecimento especial à Professora Doutora Filomena Carnide, que me ajudou muito principalmente quando a professora Vera estava de licença maternidade.

Agradecer também a minha família (pais e irmãs) por todo o suporte, tanto motivacional quanto financeiro, sem esse suporte não teria conseguido chegar até aqui. E incluo também um agradecimento super especial à minha namorada, Renata Rangel Vilela, aqueles empurrões nos momentos decisivos me deram força e motivação para conseguir concluir esse mestrado. Essa vitória dedico a vocês.

Um agradecimento gigante aos meus colegas de curso por todo companheirismo, força de vontade e por toda ajuda em todos os momentos. E um agradecimento a todos os professores que tive durante o curso, os ensinamentos de vocês me ajudaram muito. Foi um prazer trabalhar com vocês e teria sido impossível sem a vossa ajuda.

E por último, mas não menos importante, agradeço à toda equipa de investigação do laboratório de Biomecânica da FMH. Foi um prazer trabalhar com vocês e um momento de grande aprendizagem profissional.

Índice Geral

Agradecimentos	II
Índice Geral	III
Índice de tabelas	IV
Resumo	V
Abstract.....	VI
Abreviaturas.....	VII
1. Introdução	1
2. Metodologia.....	7
2.1 Desenho do estudo.....	7
2.2 Amostra	7
2.3 Avaliadores	8
2.4 Instrumentos	8
2.5 Procedimentos	10
2.6 Análise Estatística	10
3. Resultados.....	11
3.1 Fiabilidade e concordância da SF-FAB.....	11
3.2 Capacidade preditiva de quedas da SF-FAB	12
4. Discussão	14
5. Conclusão	17
6. Referências Bibliográficas.....	18
Anexo I	23

Índice de tabelas

Tabela 1. Caracterização da Amostra para a fiabilidade teste-reteste e concordância (N = 26).....	11
Tabela 2. Fiabilidade teste-reteste, concordância de cada item e do score total da SF-FAB	12
Tabela 3. Caracterização da Amostra para a capacidade preditiva da SF-FAB (N = 145)	13
Tabela 4. Regressão Logística Bivariada da SF-FAB	13
Tabela 5. Curva ROC com valores de coorte, sensibilidade e especificidade.....	14

Resumo

Objetivo: O objetivo do presente estudo foi testar a fiabilidade teste-reteste e a concordância da versão reduzida da escala “Fullerton Advanced Balance” (SF-FAB), bem como avaliar a capacidade da mesma em prever quedas na população idosa portuguesa. **Amostra:** A amostra foi constituída de 267 sujeitos, autônomos, residentes em Portugal continental e com idade ≥ 65 anos. **Métodos:** para testar a fiabilidade da escala SF-FAB foi efetuado um estudo metodológico de teste-reteste realizado por 4 avaliadores com intervalo de 2 a 7 dias entre o teste e o reteste. Para avaliar a capacidade preditiva de quedas da ferramenta foi seguido um desenho de estudo de coorte com 3 meses de duração. **Resultados:** O *score* total da escala SF-FAB mostrou boa fiabilidade teste-reteste. O valor do *Weighted Kappa* no *score* total foi de 0.89 (95% IC, 0.75 – 1.03) e a concordância foi de 50%. Os *scores* individuais (FAB4 – FAB7) mostraram boa fiabilidade. O valor do *Weighted Kappa* variou entre 0.72 – 1.00 e a concordância variou de 62 – 100%. A regressão logística binária indicou que o *score* total e os individuais da escala SF-FAB não podem ser utilizados para prever quedas (OR = 0.74 – 0.99; $p > 0.05$) na população idosa portuguesa. Não foi possível definir os valores de coorte do *score* total e dos individuais, pois a *area under curve* variou de 0.40 – 0.49 ($p > 0.05$). **Conclusão:** A escala SF-FAB mostrou-se ser uma ferramenta fiável e com boa concordância quando utilizada em idosos portugueses que vivem na comunidade. Entretanto, não deve ser utilizada como preditora de quedas pelo que não foi possível estabelecer um valor de coorte que identifique os indivíduos em risco de queda.

Palavras-chave: Idosos; Quedas; Fiabilidade teste-reteste; Capacidade preditiva; Fullerton Advanced Balance Scale.

Abstract

Purpose: The aim of this study was to evaluate the test-retest reliability and agreement of the short version of the “Fullerton Advanced Balance” scale (SF-FAB) and to determine if the scale can predict falls in the older Portuguese adults. **Participants:** The sample consisted of 267 community-dwelling older adults, independent, residing in Portugal and aged 65 or over. **Methods:** to evaluate the reliability of the SF-FAB scale a test-retest method was used; 4 raters administered the tests with an interval of 2 to 7 days between test and retest. To evaluate if the scale can predict falls, a 3 months cohort study was followed. **Results:** The total score of the SF-FAB scale showed high test-retest reliability. The Weighted Kappa of the total score was 0.89 (95% CI, 0.75 – 1.03) and the agreement was of 50%. The individual scores (FAB4 – FAB7) showed high reliability. The Weighted Kappa ranged from 0.72 – 1.00 and the agreement ranged from 62 – 100%. Binary logistic regression analysis indicated that the total and individual SF-FAB scale scores could not be used to predict falls (OR = 0.74 – 0.99; $p > 0.05$) in Portuguese elderly. It was not possible to define the cohort values of the total and individual scores of the SF-FAB scale because the area under curve ranged from 0.40 – 0.49 ($p > 0.05$). **Conclusion:** The SF-FAB scale demonstrated to be reliable and with good agreement when used in the Portuguese community-dwelling older adults. However, it should not be used as a predictor of falls because it was not possible to establish a cohort value that identify people at risk of falling.

Keywords: Elderly; Falls; Test-retest reliability; Validity; Fullerton Advanced Balance Scale.

Abreviaturas

UE: União Europeia

ADLs: Atividades da vida diária

ICC: Coeficiente de correlação intraclasse

SEM: Erro padrão de medida

MDC: Diferença mínima detetável

LOA: Limites de concordância

TUG: Timed Up and Go test

BBS: Berg Balance Scale

BESTest: Balance Evaluation Systems test

SFBBS: Short Form of Berg Balance Scale

FAB: Fullerton Advanced Balance

SF-FAB: Short form of Fullerton Advanced Balance scale

OR: Odds ratio

AUC: Area Under Curve

HFQ: Questionário de Saúde e Quedas

YPAS: Yale Physical Activity Survey

MMSE: Mini Mental State Examination

R: Reteste

IC: Intervalo de Confiança

1. Introdução

A esperança de vida está a aumentar. Segundo os dados do Eurostat, em 2017, a população da União Europeia (EU) foi estimada em 510,3 milhões de pessoas, sendo que 19,2% são pessoas idosas (com 65 anos ou mais). Isso equivale a um aumento de 0,3% em relação ao ano anterior e de 2,4% em comparação com 10 anos antes. Portugal é o quarto país da EU com a maior percentagem de pessoas idosas, 20,7%, o que representa um aumento de 0,8% em relação ao ano passado (Eurostat, 2017). Os indivíduos com idade avançada têm maior risco de desenvolver doenças crónicas e degenerativas, designadamente, doenças cardiovasculares, diabetes tipo 2, obesidade, cancro, osteoporose, osteoartrose e sarcopenia. Com o avanço da idade, há uma deterioração estrutural e funcional dos sistemas fisiológicos, que afeta tecidos e diversos órgãos. Em conjunto, essas deteriorações tem impacto na independência para executar as atividades da vida diária (ADLs) e aumentam o risco de quedas (ACSM et al., 2009; Shumway-cook et al., 2009).

A queda pode ser definida como “um evento inesperado em que o participante vem a inadvertidamente ficar no solo ou em outro nível inferior, excluindo mudanças de posição intencionais para se apoiar em móveis, paredes ou outros objetos” (Lamb, Jørstad-Stein, Hauer, & Becker, 2005). Entre 28 – 35% das pessoas idosas com idade igual ou superior a 65 anos caem uma vez por ano. Esse número aumenta para 32 – 42% para aqueles cuja idade é superior a 70 anos (World Health Organization, 2007), e essa taxa chega a aproximadamente 50% em pessoas acima dos 80 anos de idade (Todd & Skelton, 2004). Em idosos a viver na comunidade, a prevalência de sofrer pelo menos uma queda no ano é cerca de 40% (Hausdorff, Rios, & Edelberg, 2001). Um estudo realizado por Moniz-Pereira et al. (2013) encontrou resultados similares para a população idosa portuguesa. Neste estudo, que envolveu uma amostra de 1416 participantes, a prevalência de quedas (no ano anterior) foi de 38% . De acordo com o estudo realizado por Shumway-Cook et al. (2009), as quedas são um problema muito relevante experienciado pelas pessoas idosas, dada a natureza das suas consequências, designadamente, o risco incrementado de hospitalização, incapacidade, fraturas, morbidade e mortalidade. Além do risco de sofrer uma lesão grave (fratura da anca ou lesão cerebral), as pessoas que sofrem uma queda tendem a ficar ansiosas, sem confiança, e com medo de que ocorra novamente (S. H. Lee, Hee, & Kim, 2016), o que compromete a sua qualidade de vida (Öztürk, Tarsuslu, Yümin, Sertel, & Yümin, 2011).

As quedas podem ocorrer por diversos fatores, que podem ser classificados em duas categorias: Intrínsecos (específicos da pessoa) ou extrínsecos (referentes ao meio em que se encontra). São fatores de risco intrínsecos: a idade, o género, as doenças crónicas, as alterações de visão e de cognição, a redução da mobilidade e equilíbrio, os distúrbios na marcha e a fraqueza muscular. Por outro lado, existem ainda os perigos que podem fazer parte do envolvimento em que a pessoa se encontra, como: calçados não adequados, chão escorregadio ou tapetes soltos, objetos em que se pode tropeçar, móveis instáveis, falta de luz (Deandrea et al., 2010). Nesta mesma revisão sistemática, encontrou-se como fatores de risco com maior associação: o historial de quedas, problemas na marcha, uso de auxiliares na marcha, vertigem, doença de Parkinson e uso de medicamentos antiepiléticos. O padrão de marcha nas pessoas idosas tende a ser mais rígido e com menor coordenação, além de apresentarem um fraco controlo postural (Jensen, Brown, & Woollacott, 2001). As pessoas idosas são menos capazes de executar um passo rápido para evitar uma queda e demonstram dificuldades em iniciar o movimento dos braços para manter o equilíbrio quando o mesmo é alterado (McIlroy & Maki, 1996; Pavol, Runtz, Edwards, & Pai, 2002). São ainda mais suscetíveis a quedas, resultado da fraqueza nas articulações e nos músculos, além do declínio do equilíbrio (S. H. Lee et al., 2016). Adultos idosos com problemas de equilíbrio e na marcha são 3 vezes mais propensos a sofrer uma queda do que adultos idosos sem tais problemas (Hernandez & Rose, 2008). Por esta razão, ter instrumentos de terreno, devidamente validados para avaliar o equilíbrio em pessoas idosas, é um passo importante, não apenas para identificar precocemente quem está em risco de cair, mas também para planejar intervenções com maior eficácia.

Existem vários testes validados para avaliar o equilíbrio em idosos. Na sua grande maioria, essas ferramentas apresentam valores moderados a elevados de fiabilidade. De uma forma geral, a fiabilidade reflete a medida em que os resultados de determinado teste, realizado em participantes que se mantiveram estáveis, se mantêm os mesmos mediante diferentes condições (por exemplo, utilizando diferentes grupos de itens de um instrumento para testar o mesmo constructo – consistência interna – testando os participantes ao longo do tempo – teste-reteste – testando os participantes várias vezes no mesmo dia com avaliadores diferentes – inter-avaliador – ou testando os participantes em diferentes ocasiões pelo mesmo avaliador – intra-avaliador) e livre do erro de medida (i.e. o erro sistemático e aleatório que não é atribuído a verdadeira alteração no constructo que está a ser medido) (Prinsen et al., 2018). A um nível mais específico, a fiabilidade de

um instrumento é a proporção da variância total nas diferentes medições que reflete as verdadeiras diferenças entre participantes testados. Para além desta medida relativa, neste tipo de estudos, é também importante calcular-se o erro de medida.

Muitos dos estudos que avaliam a fiabilidade dos testes de terreno de equilíbrio fazem-no utilizando o coeficiente de correlação de *Pearson* ou *Spearman*. Porém esses métodos são considerados duvidosos, pois não levam em consideração os erros sistemáticos, traduzindo apenas o quanto as variáveis são correlacionadas (Prinsen et al., 2018). Para reduzir o risco de viés dos estudos de fiabilidade, alguns critérios devem ser adotados, e há uma limitação na literatura relacionada com a fiabilidade das ferramentas que avaliam equilíbrio em pessoas idosas que não levam em consideração todos esses pontos. No que diz respeito ao desenho do estudo, os participantes do estudo devem apresentar as características necessárias relacionadas ao(s) constructo(s) a ser medido no estudo e devem manter essas mesmas características em todos os momentos de avaliação. O tempo de intervalo entre o teste e o reteste deve ser longo o suficiente para o avaliador não memorizar o teste anterior e curto o suficiente para não haver a possibilidade de modificar o constructo estudado. Devem também haver condições similares em todos os momentos de avaliação. Em relação aos métodos estatísticos utilizados, para medidas contínuas, é preferível utilizar o *Intraclass Correlation Coefficient (ICC)* e o desenho teste-reteste, além de especificar o tipo de *ICC* utilizado. E para reduzir o viés do erro de medida, utiliza-se o *Erro padrão de medida (SEM)*, a *diferença mínima detetável (MDC)* ou os *limites de concordância (LOA)*. Para medidas ordinais ou quando não há normalidade dos pressupostos, o ideal é utilizar o *Weighted Cohen's Kappa*, além de especificar o peso utilizado no *weighted kappa*. Já para a redução do viés do erro de medida neste tipo de variável, como não é possível utilizar o *SEM*, *MDC* ou *LOA* a fim de encontrar a concordância, deve-se calcular a *percentagem de concordância*. E por último, evitar outras falhas que possam vir a acontecer nas avaliações. Por exemplo, as avaliações devem ter as mesmas condições e ser independentes, a primeira avaliação não pode influenciar a segunda (Prinsen et al., 2018).

Dentre os testes que avaliam apenas uma dimensão do equilíbrio, o *Timed Up & Go test (TUG)*, num estudo com uma amostra com média de idade de 65 anos que realizou uma artroplastia no joelho, apresentou excelente fiabilidade teste-reteste com $ICC(2,1) = 0,98$. O mesmo estudo apresentou ainda o *SEM* e o *MDC*, porém os testes foram realizados no mesmo dia com apenas 1 hora de intervalo entre o primeiro e segundo momento de avaliação (Yuksel, Kalkan, Cekmece, Unver, & Karatosun, 2017). No que

diz respeito a baterias de testes que focam diferentes componentes do equilíbrio, a *Berg Balance Scale (BBS)* é uma das baterias mais estudadas. Num estudo realizado com idosos que vivem em lares, a ferramenta apresentou boa fiabilidade teste-reteste com $ICC(2,1) = 0,88$. O estudo apresentou ainda o SEM e o MDC e os testes foram realizados com um intervalo entre 7 e 14 dias (Viveiro et al., 2018). Além da BBS, outras ferramentas aparecem com excelentes valores de fiabilidade. A versão norueguesa do *Balance Evaluation Systems test (BESTest)* apresentou alta fiabilidade teste-reteste (avaliador A/avaliador B = $ICC = 0,89/0,89$; $SEM = 3,9/4,3$; $MDC = 10,8/11,8$) e também a versão reduzida (mini-BESTest) apresentou boa fiabilidade teste-reteste (avaliador A/avaliador B = $ICC = 0,85/0,84$; $SEM = 1,8/1,9$; $MDC = 4,9/5,2$) quando utilizada numa amostra com idosos com risco de quedas (Hamre, Botolfsen, Tangen, & Helbostad, 2017). No entanto, a autora utilizou o ICC aleatório de uma via (1,1), sendo um método raramente utilizado (Koo & Li, 2016) e o intervalo entre os testes foi de dois dias. Já num outro estudo, também com idosos, o BESTest apresentou excelente fiabilidade teste-reteste ($ICC = 0,93$) (Wang-Hsu & Smith, 2018), entretanto a autora utilizou um intervalo maior entre um teste e outro (7 a 14 dias) e o ICC utilizado foi o aleatório de dois fatores (2,1), normalmente utilizado em estudos de fiabilidade (Prinsen et al., 2018).

Contudo, existem algumas limitações relativamente à administração destas ferramentas. No caso da BBS e do BESTest o tempo de aplicação das baterias é muito longo. O TUG, apesar de ser de rápida aplicação, é pouco desafiante para pessoas com boa capacidade funcional e de equilíbrio, não sendo capaz de detetar problemas mais específicos de equilíbrio (Mackintosh, Datson, & Fryer, 2006). A BBS aparece como o teste de terreno mais utilizado para avaliar o equilíbrio e risco de queda em idosos incapacitados e fragilizados. No entanto, esta ferramenta mostrou-se muito menos precisa para a avaliação do risco de queda em pessoas idosas com boa capacidade funcional e sem problemas de equilíbrio (Boulgarides, McGinty, Willett, & Barnes, 2003; Brauer, Burns, & Galley, 2000). Para além disso, a BBS consiste em 14 itens com critérios de pontuação variados em todos eles, o que se torna um problema para quem não tem muita prática na aplicação deste teste e apresenta uma consistência interna extremamente alta, o que indica redundância nos itens (Chou et al., 2006). A fim de tornar a BBS mais rápida e simples de administrar, Chou et al. (2006) desenvolveu uma forma simplificada da ferramenta em pessoas com AVC, a *Short Form of Berg Balance Scale (SFBBS)*, com valores psicometricamente similares à versão original. Karthikeyan, Sheikh, & Chippala (2012) encontrou excelente fiabilidade teste-reteste da SFBBS em idosos ($ICC = 0,95$).

Porém, este estudo apresentou algumas limitações no que diz respeito à metodologia adotada para estudar fiabilidade da ferramenta, nomeadamente, não foi definido o tipo de ICC utilizado, e não foi calculado o SEM e nem o MDC.

Para minimizar os pontos negativos das ferramentas mencionadas, foi desenvolvida a *Fullerton Advanced Balance Scale (FAB)* (Rose, Lucchese, & Wiersma, 2006), uma ferramenta criada para avaliar, em terreno, as várias dimensões do equilíbrio, tanto em situações estáticas quanto dinâmicas, em idosos com independência funcional e que vivem em comunidade. A FAB é, ainda, fácil e rápida de administrar (demora de 10 a 12 minutos para ser realizada). As autoras originais reportaram uma elevada fiabilidade para esta bateria de testes. No entanto, neste estudo a fiabilidade foi testada através do coeficiente de correlação de Spearman (*Spearman's rank correlation coefficients* (p) = 0,96), um método que não é o ideal por não ter em conta o erro sistemático. A partir da versão original, foram estudadas outras versões da FAB e elas apresentam ótimos resultados, porém utilizando o mesmo método utilizado pelas autoras originais (Iyigun et al., 2018; Schott, 2011). Já a FAB persa obteve excelente fiabilidade teste-reteste (ICC = 0,97) (Azad, Sabet, Taghizadeh, & Mohammadi-Nezhad, 2018). Contudo, este estudo apresentou algumas limitações no que diz respeito a fiabilidade da ferramenta: 1. não foi definido o tipo de ICC utilizado; 2. foi utilizado o coeficiente de *kappa* para medir a concordância e não a fiabilidade de medidas ordinais. Comparativamente à SFBBS, a fim de reduzir ainda mais o tempo de avaliação, Hernandez & Rose (2008) estabeleceram uma versão reduzida da FAB, denominada SF-FAB. Essa versão é composta pelos 4 itens com maior capacidade preditora de quedas da versão original (itens 4-7), desenhada para reduzir o tempo total do teste. E, apesar de as autoras não disponibilizarem a fiabilidade teste-reteste de cada item da FAB, sabe-se que as correlações destes itens foram significativos e apresentaram boa fiabilidade (Rose et al., 2006)

Afim de ser capaz de detetar pessoas com risco de queda, uma ferramenta além de ser fiável deve ser válida para tal, por isso, relativamente à capacidade em prever quedas, há na literatura alguns estudos que analisam as capacidades preditivas da FAB. Devido ao facto de a BBS ser uma das ferramentas mais estudadas nesta área, Jeon & Kim (2017) comparou a capacidade preditiva das duas ferramentas em idosos que vivem em comunidade, e as duas ferramentas apresentaram boas capacidades preditivas. A FAB apresentou significância no *odds ratio (OR)* da regressão logística bivariada ($p < 0,05$) com valor de 1,22. O *cut-off* foi de 22 – 40 com a *area under curve (AUC)* = 0,72 e a sensibilidade e especificidade iguais a 0,85 e 0,65, respetivamente. Já a BBS apresentou

resultados muito próximos, com $OR = 1,3$ ($p < 0,05$), o *cut-off* de 40 – 56, $AUC = 0,70$ e sensibilidade e especificidade iguais a 0,82 e 0,67, ou seja, para cada aumento de 1 ponto no *score* total da FAB e BBS, aumentou em 22% e 30% a probabilidade de o sujeito não cair (*non-faller*). O primeiro estudo que avaliou as propriedades preditivas da FAB, realizado por Hernandez & Rose (2008) teve um $OR = 0,902$ ($p < 0,001$), o *cut-off* foi de 25 – 40 e a sensibilidade e especificidade foram de 74,6% e 52,6%, respetivamente, ou seja, a cada redução de 1 ponto no *score* total da FAB a probabilidade de quedas aumentou em 8%. No entanto, a autora não disponibilizou a AUC, responsável por indicar se a ferramenta é de facto preditiva ($AUC > 0,50$) ou não ($AUC \leq 0,50$) (Park, 2013). No estudo realizado por Azad et al. (2018) a regressão logística bivariada foi significativa ($p < 0,001$) e o *cut-off* foi de 32 – 40 com sensibilidade e especificidade iguais a 81,25% e 52,78%. Entretanto, o autor não especificou a AUC e nem o valor do OR.

Em todos os estudos acerca da capacidade preditiva da FAB foram utilizados, para o historial de quedas, o auto-relato retrospectivo de 6 ou 12 meses, o que exige uma boa memória dos participantes, podendo aumentar o viés do estudo. Por esta razão, Hernandez & Rose (2008) sugerem que seja utilizado um estudo prospetivo, para que a incidência de quedas seja monitorada durante um período de tempo (1 ano, por exemplo), tendo uma avaliação mais precisa do historial de quedas. Os mesmo estudos também utilizaram, antes de calcular a regressão logística bivariada, o teste de *Hosmer and Lemeshow* para avaliar a qualidade de ajuste do modelo. No entanto, este teste é indicado para uma amostra com tamanho de 400 sujeitos, o que não é o caso dos estudos. Os autores deveriam ter utilizado o teste *Qui-quadrado de Pearson* (X^2) visto que é similar ao *Hosmer and Lemeshow* para avaliar a qualidade de ajuste do modelo (Bewick, Cheek, & Ball, 2005).

Considerando a importância de ter ferramentas fiáveis e com boa capacidade em prever quedas para os profissionais de exercício e saúde que intervêm com a população idosa, o presente estudo tem como objetivos: (1) avaliar a fiabilidade teste-reteste e a concordância da SF-FAB e (2) avaliar a capacidade da mesma ferramenta em prever quedas na população idosa portuguesa.

2. Metodologia

2.1 Desenho do estudo

Para testar a fiabilidade da bateria SF-FAB foi efetuado um estudo metodológico de teste-reteste. Para a avaliação da capacidade preditiva dos itens FAB 4 a FAB 7, na predição das quedas, foi desenvolvido um estudo de coorte prospetivo com 3 meses de seguimento.

2.2 Amostra

A amostra foi constituída por uma coorte de 267 participantes do projeto de investigação “Frailty assessment in elderly people and risk stratification for falls and disability: determination of an objective and simplified frailty index” (ref: EXPL/DTP/DES/1915/2013). Para o recrutamento, foram convidados utentes dos centros de dia dos municípios de Oeiras, Lisboa e Loures. Foram considerados os seguintes critérios de inclusão: ter idade igual ou superior a 65 anos, ser autónomo, compreender e falar corretamente a língua portuguesa, ser capaz de andar independentemente. Como critérios de exclusão foram definidos: a presença de doença cerebrovascular ou cognitiva (Mini Mental State Examination < 24) (Karthikeyan et al., 2012), ter mobilidade com auxiliares de marcha, ter prótese na coluna vertebral ou qualquer segmento do membro inferior, ou estar em recuperação de doença aguda.

Para o cálculo da amostra do estudo da fiabilidade teste-reteste foi utilizada a equação de Kraemer & Blasey (2016). Considerou-se o valor mínimo do ICC (0,7) e o valor ideal (0,9), um nível de significância de 5% e uma potência de 80%, obtendo-se um mínimo de 18 participantes para o estudo. No entanto, para aumentar a potência dos resultados e de forma a ter em conta possíveis desistências e/ou exclusões, 30 participantes foram convidados a realizarem uma 2ª avaliação (reteste).

Para a avaliação da capacidade preditiva da SF-FAB considerou-se a prevalência de quedas em idosos portugueses de 38% (Moniz-Pereira et al., 2013), e considerando os mesmos níveis de significância e de potência da amostra, foi obtido um número mínimo de 118 participantes (Hsieh, Block, & Larsen, 1998). No entanto, para aumentar a

potência dos resultados e de forma a ter em conta possíveis desistências e/ou exclusões, utilizou-se toda a coorte previamente mencionada.

Todos os participantes leram e assinaram o consentimento informado. O protocolo do estudo foi avaliado e aprovado pelo Conselho de Ética para a Investigação da Faculdade de Motricidade Humana, da Universidade de Lisboa.

2.3 Avaliadores

A recolha de dados foi realizada por 4 avaliadores (com formação académica em ciências do desporto – especialidade em exercício e saúde –, ergonomia, fisioterapia e nutrição), treinados especificamente pela equipa de investigação para a aplicação dos diversos instrumentos que suportam a avaliação dos fatores preditores e do *outcome*. As sessões de formação totalizaram 15 horas, distribuídas em 5h teóricas, subordinadas aos procedimentos dos testes, bem como à compreensão dos constructos, e 10 horas práticas, sendo 5 horas em sala de aula, com recursos de vídeo e 5 horas de avaliação no terreno, envolvendo pessoas idosas que, voluntariamente, se disponibilizaram para participar na formação. Em complemento à formação, cada avaliador, recebeu um manual de procedimentos e guiões de instruções de aplicação dos diversos instrumentos.

A fim de garantir a qualidade dos avaliadores no que diz respeito às avaliações do equilíbrio, foi realizada a fiabilidade inter-observador e obteve-se excelente fiabilidade do *score* total da SF-FAB, com o valor do *Weighted Cohen's Kappa* peso quadrático de 0,90 ($p < 0,001$) em um intervalo de confiança de 0,81 – 1,00.

2.4 Instrumentos

Na primeira avaliação foram aplicados o questionário de saúde e quedas (HFQ), o *Yale Physical Activity Survey* (YPAS), o *Mini mental State Examination* (MMSE) e foram realizadas as baterias do *Senior Fitness Test* e a SF-FAB. O HFQ foi aplicado afim de obter dados sociodemográficos, de saúde, condição visual e auditiva, histórico médico (consultas médicas, hospitalizações e cirurgias), medo de cair, capacidade de realizar atividades do dia-a-dia após a queda, prevalência de quedas (no último ano) e características da queda (local, circunstâncias e consequências das últimas 3 quedas reportadas). Estas variáveis permitiram avaliar o *outcome* principal do estudo e uma queda foi definida como “um evento inesperado em que o participante vem a

inadvertidamente ficar no solo ou em outro nível inferior, excluindo mudanças de posição intencionais para se apoiar em móveis, paredes ou outros objetos”(Lamb et al., 2005).

A atividade física foi avaliada através do YPAS, um questionário com perguntas acerca de uma típica semana durante o mês para avaliar a intensidade, frequência e duração em 5 dimensões distintas de atividade física: 1) atividade vigorosa; 2) caminhada; 3) movimento; 4) tempo de pé e 5) tempo sentado. Cada *score* parcial é multiplicado pelo fator de ponderação (compreendido entre 5 e 1, respetivamente) a fim de calcular os índices individuais, que somados permitem determinar o índice global de atividade física (Dipietro, Caspersen, Ostfeld, & Nadel, 1993).

O MMSE foi utilizado para avaliar a capacidade cognitiva ou presença de doença cerebrovascular dos participantes, e foi utilizado como critério de exclusão do presente estudo. O teste é dividido em duas sessões, a primeira que consiste em respostas de perguntas, capacidade de memorizar e atenção; o score máximo é 21. A segunda parte testa a habilidade de nomear objetos, seguir comandos verbais e escritos, escrever uma frase espontaneamente, e copiar o desenho de polígonos; o score máximo é 9. O teste não é temporizado e o score máximo é 30 (Folstein, Folstein, & McHugh, 1975).

A avaliação da performance funcional foi realizada a partir dos parâmetros de força e equilíbrio. A força, potência e mobilidade dos membros inferiores foram medidos através do “2-minute walk test”, “30 sec Chair-Stand test” e “Timed up-and-go test” da bateria *Senior Fitness Test* (Rikli & Jones, 1999). Já a avaliação do equilíbrio foi realizada através da aplicação dos itens 4-7 da escala FAB, baseado na sua fiabilidade como teste clínico e de terreno (Hernandez & Rose, 2008). Os dois primeiros itens da escala FAB avaliam o equilíbrio dinâmico e os dois últimos, o equilíbrio estático. O item 4, consiste na transposição de um banco de 15cm, com apenas um apoio sobre o banco. O item 5, consiste na realização de 10 passos em linha reta sobre uma linha colocada no chão. No item 6, é pedido ao sujeito que permaneça em pé, sobre apenas um apoio e com os braços cruzados sobre o peito até no máximo 20 segundos. Por fim, o item 7, consiste em permanecer em pé sobre uma superfície de esponja com os olhos fechados e braços cruzados sobre o peito durante um período máximo de 20 segundos (Rose et al., 2006).

2.5 Procedimentos

Para a avaliação da fiabilidade teste-reteste, os participantes foram avaliados em duas ocasiões diferentes pelo mesmo avaliador, com um intervalo mínimo de 2 dias e máximo de 7 dias (Karthikeyan et al., 2012; Rose et al., 2006) (Anexo I).

Para a determinação do valor preditivo do equilíbrio na ocorrência de quedas, os participantes foram reavaliados 3 meses após o *baseline*. A duração do período de *follow-up* foi determinada de acordo com as condicionantes do projeto de investigação “Frailty assessment in elderly people and risk stratification for falls and disability: determination of an objective and simplified frailty index” (ref: EXPL/DTP/DES/1915/2013). Os participantes foram divididos em “sem historial de quedas”, os que não caíram nos 3 meses, e “com historial de quedas”, os que caíram uma ou mais vezes nesse mesmo período.

2.6 Análise Estatística

A análise estatística foi realizada a partir do *software* SPSS versão 25.0 (IBM, Armonk, NY), para um nível de significância de $p < 0,05$. Foi utilizada a estatística descritiva para a caracterização das variáveis demográficas, de saúde e de equilíbrio. Após a realização dos pressupostos, foram escolhidos testes não-paramétricos pelo facto de a amostra não ter uma distribuição normal.

A fiabilidade de cada um dos testes individualmente e da pontuação total da SF-FAB foi calculada através do *Weighted Cohen's Kappa* com um esquema de peso quadrático, por se tratarem de medidas ordinais com mais de 2 categorias. Quando utilizado o peso quadrático, o teste é idêntico ao *ICC* concordância. Foi considerado como aceitável um valor $\geq 0,70$ (Prinsen et al., 2018). Para a concordância foi utilizado a percentagem de concordância entre os dois momentos de avaliação, calculado no Microsoft Excel 2016 (Microsoft Corp., Redmond, WA). Foi calculada também a percentagem das diferenças entre pontuações que se verificou ser positiva e negativa, para verificar se as pontuações atribuídas no segundo momento de avaliação foram predominantemente superiores ou inferiores em relação ao primeiro momento de avaliação.

As propriedades discriminativas da escala SF-FAB no que respeita ao *outcome* quedas, foi determinada a partir da dicotomização da variável quedas (0 = sem historial de quedas no *follow-up* de 3 meses e 1 = com historial de quedas no *follow-up* de 3 meses) e foi

utilizada a regressão logística bivariada, tendo sido utilizado o *odds ratio* (OR) como estimador da probabilidade de predição de quedas e foi considerado um nível de significância de p inferior a 0,05. Para assegurar a qualidade de ajuste do modelo foi utilizado o teste *Qui-quadrado de Pearson* (X^2) já que não era possível utilizar o teste de *Hosmer and Lemeshow* devido ao tamanho da amostra ser inferior a 400 (Bewick et al., 2005). A capacidade preditiva da escala SF-FAB foi determinada pela *area under curve* (AUC). Foi considerado como aceitável um valor $> 0,5$ e quanto mais próximo a 1,0 maior o valor preditivo da ferramenta (Park, 2013) e os *cut-offs* dos scores individuais e total foram determinados quando o somatório da sensibilidade e especificidade se aproximavam a 1 (Bewick, Cheek, & Ball, 2004).

3. Resultados

3.1 Fiabilidade e concordância da SF-FAB

Dos 30 idosos que realizaram o teste-reteste da SF-FAB, 4 foram excluídos por ter um intervalo maior do que 7 dias entre uma avaliação e outra. As características dos 26 idosos são mostradas na tabela 1.

Tabela 1. Caracterização da Amostra para a fiabilidade teste-reteste e concordância (N = 26)

	N (%)	Média±DP (Me)	Amplitude
<i>Masculino</i>	12 (46,2)	-	-
<i>Feminino</i>	14 (53,8)	-	-
<i>Idade (anos)</i>	-	72,3±6,0 (72,0)	65 - 84
<i>Quedas baseline</i>	6 (23)	0,32±0,69	0 - 3
<i>Atividade_vigorosa</i>	-	2,31±5,33 (0,00)	-
<i>Atividade caminhada</i>	-	10,00±8,56 (8,00)	-
<i>Movimento</i>	-	8,73±3,01 (9,00)	-
<i>Massa corporal (kg)</i>	-	71,59±12,66 (68,00)	54 - 104
<i>Altura (m)</i>	-	1,62±0,09 (1,59)	1,45 - 1,78
<i>FAB4</i>	-	3,85±0,78 (4,00)	0 - 4
<i>FAB5</i>	-	2,88±1,24 (3,00)	0 - 4
<i>FAB6</i>	-	2,92±1,32 (3,50)	0 - 4
<i>FAB7</i>	-	3,46±0,98 (4,00)	0 - 4
<i>SCORE TOTAL SF-FAB</i>	-	13,12±3,51 (14,00)	0 - 16
<i>R FAB4</i>	-	3,85±0,78 (4,00)	0 - 4
<i>R FAB5</i>	-	2,96±1,11 (3,00)	0 - 4
<i>R FAB6</i>	-	2,92±1,29 (3,50)	0 - 4
<i>R FAB7</i>	-	3,46±1,02 (4,00)	0 - 4
<i>R_SCORE_TOTAL_SF-FAB</i>	-	13,19±3,54 (14,00)	0 - 16

FAB: Fullerton Advanced Based Scale; R: reteste; DP: desvio padrão; Me: mediana

A tabela 2 mostra a fiabilidade, a percentagem de concordância e a diferença entre *scores* positiva e negativa dos testes individuais e do *score* total da SF-FAB. A ferramenta obteve boa fiabilidade, o *weighted kappa* do *score* total foi 0,89 e o intervalo de confiança de 0,75 – 1,03. Todos os testes, individualmente, tiveram boa fiabilidade e variaram 0,72 – 1,00. Destaca-se a FAB4 (*weighted kappa* = 1,00) com excelente fiabilidade e a FAB5 (*weighted kappa* = 0,85). A percentagem de concordância da SF-FAB variou de 50 – 100%. A FAB4 obteve 100% de concordância entre o primeiro e segundo momento de avaliação e a FAB7 obteve 80%. A diferença entre *scores* positiva variou de 12 – 27% enquanto a negativa de 8 – 23%. Não houve uma diferença muito grande entre quem errou por proporcionar um *score* maior ou menor na segunda avaliação em comparação com a primeira.

Tabela 2. Fiabilidade teste-reteste, concordância de cada item e do *score* total da SF-FAB

	Weighted Cohen's Kappa (IC95%)	Z	p	Concordância (%)	Diferença entre scores positiva (%)	Diferença entre scores negativa (%)
<i>FAB4</i>	1,00 (1,00 -1,00)	5,09	< 0,001	100	-	-
<i>FAB5</i>	0,85 (0,73 – 0,98)	4,40	< 0,001	73	15	12
<i>FAB6</i>	0,74 (0,54 – 0,93)	3,79	< 0,001	62	19	19
<i>FAB7</i>	0,72 (0,37 – 1,07)	3,70	< 0,001	80	12	8
<i>SCORE TOTAL SF-FAB</i>	0,89 (0,75 – 1,03)	4,54	< 0,001	50	27	23

Z: estatística do teste; FAB: Fullerton Advanced Based Scale; SF-FAB: Short form of Fullerton Advanced Based Scale

3.2 Capacidade preditiva de quedas da SF-FAB

Dos 267 idosos da amostra inicial, 22 foram excluídos por viverem em casa ou clube de repouso, 83 por apresentarem *score* no *Mini Mental* abaixo de 24, 11 por não serem autónomos nas tarefas do dia-a-dia e 6 por utilizarem auxiliares de marcha para tarefas locomotoras. As características dos 145 idosos são apresentadas na tabela 3.

Tabela 3. Caracterização da Amostra para a capacidade preditiva da SF-FAB (N = 145)

	N (%)	Média±DP (Me)	Amplitude
<i>Masculino</i>	55 (37,9)	-	-
<i>Feminino</i>	90 (62,1)	-	-
<i>Idade</i>	-	74,8±6,4 (75,0)	65 - 91
<i>Estado civil (casado)</i>	74 (51)	-	-
<i>Escolaridade (básico)</i>	61 (42)	-	-
<i>Vive acompanhado</i>	89 (61,4)	-	-
<i>Vive em casa própria</i>	143 (98,6)	-	-
<i>Quedas baseline</i>	42 (29)	-	-
<i>Medo de cair (nunca)</i>	96 (66,2)	-	-
<i>Atividade vigorosa</i>	-	4,69±10,70 (0,00)	-
<i>Atividade caminhada</i>	-	14,32±12,73 (16,00)	-
<i>Movimento</i>	-	8,75±2,90 (9,00)	-
<i>Massa corporal (kg)</i>	-	69,26±12,18 (68,00)	45 - 102
<i>Altura (m)</i>	-	1,58±0,90 (1,58)	1,31 - 1,82
<i>FAB4</i>	-	3,72±0,83 (4,00)	0 - 4
<i>FAB5</i>	-	2,94±1,16 (3,00)	0 - 4
<i>FAB6</i>	-	2,76±1,27 (3,00)	0 - 4
<i>FAB7</i>	-	3,46±1,00 (4,00)	0 - 4
<i>SCORE TOTAL SF-FAB</i>	-	12,96±3,53 (14,00)	0 - 16
<i>Quedas nos últimos 3 meses</i>	13 (9)	-	0 - 3

FAB: Fullerton Advanced Based Scale; DP: desvio padrão; Me: mediana

A tabela 4 mostra os resultados da *regressão logística binária* do *score* total da SF-FAB e dos testes individuais em relação ao *follow-up* de 3 meses. O OR da SF-FAB variou de 0,74 – 0,99, não sendo encontrados resultados estatisticamente significativos. Sendo assim, não pode ser considerada uma ferramenta capaz de prever quedas em idosos portugueses em 3 meses de *follow-up*. Os valores de qui-quadrado (X^2) variaram de 1,92 – 10,52 e os valores de p, tanto do *score* total quanto dos *scores* individuais, obtiveram valores maiores do que 0,05, assegurando a qualidade de ajuste do modelo.

Tabela 4. Regressão Logística Bivariada da SF-FAB

	OR (IC95%)	p	X^2	p
<i>FAB4</i>	0,74 (0,44 – 1,24)	0,26	6,72	0,15
<i>FAB5</i>	0,87 (0,55 – 1,39)	0,58	3,27	0,51
<i>FAB6</i>	0,90 (0,58 – 1,41)	0,67	4,49	0,34
<i>FAB7</i>	0,99 (0,56 – 1,77)	0,99	1,92	0,74
<i>SCORE_TOTAL_SF-FAB</i>	0,95 (0,82 – 1,09)	0,48	10,52	0,65

OR: Odds ratio; X^2 : Qui-quadrado de Pearson; FAB: Fullerton Advanced Based Scale

A tabela 5 mostra o valor de corte do *score* total e de cada teste individual capaz de prever quedas em idosos portugueses. A AUC variou de 0,40 – 0,49, não sendo encontrados resultados estatisticamente significativos para esses valores. Por esta razão, não foi possível definir qual o valor de cada teste e do *score* total da SF-FAB, ou seja, o valor capaz de determinar se o idosos tem mais ou menos risco de sofrer uma queda.

Tabela 5. Curva ROC com valores de corte, sensibilidade e especificidade

	AUC (IC95%)	p	Cut-off score	Especificidade	Sensibilidade
<i>FAB4</i>	0,40 (0,23 – 0,58)	0,28	3,50	87	69
<i>FAB5</i>	0,49 (0,31 – 0,66)	0,91	2,50	75	76
<i>FAB6</i>	0,46 (0,30 – 0,61)	0,64	2,50	56	61
<i>FAB7</i>	0,46 (0,30 – 0,62)	0,65	3,50	71	61
<i>SCORE TOTAL</i>	0,42 (0,25 – 0,59)	0,37	12,50	65	53
<i>SF-FAB</i>					

AUC: Area under curve; FAB: Fullerton Advanced Based Scale

4. Discussão

Os objetivos do presente estudo foram: (1) avaliar a fiabilidade teste-reteste e a percentagem da concordância da SF-FAB; e, (2) avaliar a capacidade da mesma ferramenta em prever quedas na população idosa portuguesa. Os resultados mostram que a SF-FAB é uma ferramenta com boa fiabilidade teste-reteste em idosos portugueses. No que diz respeito às suas propriedades discriminativas em relação as quedas, e ao contrário do que seria de esperar, a SF-FAB não se mostrou uma ferramenta preditiva de quedas em portugueses em um *follow-up* de 3 meses.

Esse é o primeiro estudo que avalia a fiabilidade e a concordância da SF-FAB. Este estudo encontrou boa fiabilidade do *score* total desta ferramenta sendo, no entanto, de difícil comparação com os resultados obtidos com versões mais atuais da ferramenta. A versão turca (Iyigun et al., 2018) obteve resultados similares ao presente estudo, porém o autor utilizou o coeficiente de correlação de *Spearman*, o que não é um bom método para a avaliação da fiabilidade, pois não leva em consideração os erros sistemáticos (Prinsen et al., 2018). Já a versão persa (Azad et al., 2018), que também obteve excelente fiabilidade do *score* total, não especificou o tipo de ICC que foi utilizado no estudo não sendo possível, assim, saber se o tipo de ICC utilizado era o ideal. Para o presente estudo

foi utilizado o *weighted kappa* peso quadrático devido à falta de normalidade dos pressupostos.

Embora não existam outros estudos de fiabilidade da versão reduzida da FAB, resultados similares foram obtidos em versões reduzidas de outras ferramentas que avaliam o equilíbrio. Num estudo com idosos, a versão reduzida da BBS (SFBBBS) obteve excelente fiabilidade teste-reteste do *score* total, porém o autor não especificou o tipo de ICC utilizado (Karthikeyan et al., 2012). Comparativamente a versão reduzida do BESTest (mini-BESTest), também com idosos em risco de quedas, os resultados foram similares. A ferramenta também apresentou boa fiabilidade e todos os pontos para reduzir o viés foram seguidos. No caso deste estudo, foi possível ainda calcular o SEM e o MDC pois utilizaram o ICC para a fiabilidade (Hamre et al., 2017).

O nosso estudo também encontrou boa fiabilidade em cada item da SF-FAB. Num estudo realizado por Azad et al. (2018) resultados similares foram encontrados na versão persa da FAB. No entanto, o autor utilizou o ICC para a fiabilidade e o *Cohen's kappa* para a concordância dos itens da ferramenta. De facto, *kappa* é uma medida de fiabilidade e não de concordância. Para a concordância de medidas ordinais utiliza-se a percentagem de concordância (Prinsen et al., 2018). A percentagem de concordância do presente estudo foi boa (> 70%), com exceção da FAB6 (62%). Ao analisar os dados verificou-se que grande parte da concordância ocorreu em quem obteve *score* máximo ou quem obteve zero nos testes e a maior parte da discordância verificou-se em pessoas que obtiveram *scores* intermédios nos testes. Isso mostra que fica mais claro para o avaliador analisar a performance de pessoas com elevada ou muito baixa valência física (equilíbrio) do que aquelas pessoas que não estão em nenhum dos extremos. Utilizando a FAB6 como exemplo, fica muito mais claro avaliar pessoas que não conseguem manter o equilíbrio sobre um apoio ou que conseguem facilmente o fazer por 20 segundos, do que avaliar pessoas que variam entre 5, 12, 15 segundos nesta mesma posição (Rose et al., 2006). Por outro lado, o que poderá explicar a discordância de quem obteve *scores* intermédios, é a variabilidade natural dos participantes de uma sessão para outra.

No que diz respeito a diferença dos *scores* ser positiva ou negativa, os resultados deste estudo foram muito próximos, com valores positivos levemente maiores do que os negativos, o que implica que alguns participantes tiveram maior *score* no segundo momento de avaliação. Desta forma, fica claro não ter havido efeito do treino nos participantes ou efeito da memória dos avaliadores e não houve mudança do atributo a ser avaliado (Sim & Wright, 2005); caso contrário a maior parte das diferenças seriam

positivas no segundo momento da avaliação. Um ponto positivo deste estudo foi a utilização de 2 a 7 dias de intervalo entre os testes. Sim & Wright (2005) estabeleceram que o intervalo entre 2 e 14 dias é o normal, mas depende do constructo a ser medido. Além deste intervalo estar de acordo com o que se encontra na literatura, é um intervalo de tempo semelhante ao utilizado em outros estudos que avaliaram a fiabilidade em versões reduzidas de outras baterias de equilíbrio (Karthikeyan et al., 2012; Langley & Mackintosh, 2007). Porém, uma vez que não há na literatura definição do intervalo ideal entre o teste e o reteste em estudo de fiabilidade de instrumentos que avaliam o equilíbrio e que esta qualidade física parece poder ser uma das primeiras a diminuir durante o processo de destreino (M. Lee, Lim, Lee, Kim, & Yoon, 2017), seria interessante realizar um estudo com essas características para reduzir a possibilidade de viés de futuros estudos.

Apesar de não ter havido resultados estatisticamente significativos do OR tanto do *score* total quanto dos *scores* individuais da SF-FAB, destaca-se o facto de ter sido utilizada uma metodologia prospectiva, no que diz respeito ao *follow-up* das quedas nos últimos 3 meses de cada sujeito da amostra, permitindo o controlo de diferentes fontes de viés e uma avaliação mais precisa das propriedades preditivas da SF-FAB (Hernandez & Rose, 2008). Todos os outros estudos utilizaram um desenho retrospectivo, obtendo assim melhores resultados que o encontrado neste estudo (Azad et al., 2018; Hernandez & Rose, 2008; Jeon & Kim, 2017). Ao analisarmos as características das amostras utilizadas nestes estudos, verifica-se que essa diferença entre o resultado obtido no presente estudo quando comparado com as outras versões da FAB não têm relação com a caracterização da amostra, uma vez que a média de idade (75 anos), o nível de atividade física (5 dias por semana) e o género (60% sexo feminino) são semelhantes em todos os estudos. O mesmo acontece com o facto da AUC tanto do *score* total quanto dos *scores* individuais ficarem abaixo de 0,5 e sem significância, impossibilitando assim definir o *cut-off* que distingue um caidor (*faller*) de um não caidor (*non-faller*). Esse seria o primeiro estudo que encontraria a capacidade preditiva da SF-FAB, uma vez que só há na literatura o estudo do Jeon & Kim (2017), que leva em consideração o resultado da AUC, que encontrou o *cut-off* de 22 dos 40 pontos totais da versão completa da FAB e o OR de 1,22, o que significa que para cada aumento de 1 ponto no *score* total da FAB, aumenta em 22% a chance de o sujeito não cair (*non-faller*).

O presente estudo apresenta também algumas limitações. O facto do *score* total não cumprir os pressupostos para utilizar os testes paramétricos, não nos permitiu calcular

o SEM e a MDC. Estas medidas seriam muito importantes para os profissionais no terreno conhecerem qual o valor mínimo do *score* que seria de facto uma melhoria e não um erro de medida. Ademais, o facto do período de *follow-up* ser curto, resultando num número reduzido de indivíduos que caíram (apenas 9%), poderá ter influenciado os resultados da capacidade da SF-FAB em prever quedas. No entanto, no contexto deste projeto, não foi possível manter os recursos humanos necessários para continuar o seguimento.

Futuramente, seria útil investigar a fiabilidade inter-observador da SF-FAB, assim como a capacidade da ferramenta prever quedas em idosos utilizando um follow-up com 12 meses, com avaliações intercalares de 3 meses. Acrescenta-se ainda a sugestão de realização de um estudo que avalie o poder de resposta da ferramenta, relativamente a programas de exercício.

5. Conclusão

Os resultados deste estudo mostram que a SF-FAB apresenta boa fiabilidade teste-reteste em idosos portugueses que vivem na comunidade. A ferramenta apresentou boa percentagem de concordância, principalmente nos participantes que obtiveram máxima ou mínima pontuação.

Este estudo também mostra que a SF-FAB não deve ser utilizada como preditora de quedas em um *follow-up* de 3 meses em idosos portugueses que vivem na comunidade, pelo que não foi possível estabelecer um valor de corte que identifique os indivíduos em risco de queda

6. Referências Bibliográficas

- ACSM, Chodzko-Zajko, W. J., Proctor, D. N., Fiatarone Singh, M. A., Minson, C. T., Nigg, C. R., ... Skinner, J. S. (2009). Exercise and physical activity for older adults. *Medicine and Science in Sports and Exercise*, 41(7), 1510–1530. <https://doi.org/10.1249/MSS.0b013e3181a0c95c>
- Azad, A., Sabet, A., Taghizadeh, G., & Mohammadi-Nezhad, T. (2018). Clinical assessment of Persian translation of Fullerton Advanced Balance Scale in community-dwelling older adults. *Disability and Rehabilitation*, 0(0), 1–7. <https://doi.org/10.1080/09638288.2018.1503731>
- Bewick, V., Cheek, L., & Ball, J. (2004). Statistics review 13: Receiver operating characteristics curves. *Critical Care*, 8(6), 508–512. <https://doi.org/10.1186/cc3000>
- Bewick, V., Cheek, L., & Ball, J. (2005). Statistics review 14: Logistic regression. *Critical Care*, 9(1), 112–118. <https://doi.org/10.1186/cc3045>
- Boulgarides, L. K., McGinty, S. M., Willett, J. a, & Barnes, C. W. (2003). Use of clinical and impairment-based tests to predict falls by community-dwelling older adults. *Physical Therapy*, 83(4), 328–39. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12665404>
- Brauer, S. G., Burns, Y. R., & Galley, P. (2000). A Prospective Study of Laboratory and Clinical Measures of Postural Stability to Predict. *Journal of Gerontology*, 55(8), 469–476.
- Chou, C. Y., Chien, C. W., Hsueh, I. P., Sheu, C. F., Wang, C. H., & Hsieh, C. L. (2006). Developing a short form of the Berg Balance Scale for people with stroke. *Phys Ther*, 86(2), 195–204. <https://doi.org/10.1644/5333>
- Deandrea, S., Lucenteforte, E., Bravi, F., Foschi, R., La Vecchia, C., & Negri, E. (2010). Risk Factors for Falls in Community-dwelling Older People. *Epidemiology*, 21(5), 658–668. <https://doi.org/10.1097/EDE.0b013e3181e89905>
- Dipietro, L., Caspersen, C. J., Ostfeld, A. M., & Nadel, E. R. (1993). A survey for assessing physical activity among older adults. *Medicine and Science in Sports and Exercise*, 25(5), 628–42. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8492692>
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”. A

- practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–98. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)
- Hamre, C., Botolfsen, P., Tangen, G. G., & Helbostad, J. L. (2017). Interrater and test-retest reliability and validity of the Norwegian version of the BESTest and mini-BESTest in people with increased risk of falling. *BMC Geriatrics*, 17(1), 92. <https://doi.org/10.1186/s12877-017-0480-x>
- Hausdorff, J. M., Rios, D. A., & Edelberg, H. K. (2001). Gait variability and fall risk in community-living older adults: A 1-year prospective study. *Archives of Physical Medicine and Rehabilitation*, 82(8), 1050–1056. <https://doi.org/10.1053/apmr.2001.24893>
- Hernandez, D., & Rose, D. J. (2008). Predicting Which Older Adults Will or Will Not Fall Using the Fullerton Advanced Balance Scale. *Archives of Physical Medicine and Rehabilitation*, 89(12), 2309–2315. <https://doi.org/10.1016/j.apmr.2008.05.020>
- Hsieh, F. Y., Block, D. A., & Larsen, M. D. (1998). A Simple Method of Sample Size Calculation for Linear and Logistic Regression. *Statistics in Medicine*, 17(February 1997), 1623–1634. Retrieved from <http://www.zi-berlin.de/morbilitaetsanalyse/downloads/ADT-Panel-version3.pdf>
- Iyigun, G., Kirmizigil, B., Angin, E., Oksuz, S., Can, F., Eker, L., & Rose, D. J. (2018). The reliability and validity of the Turkish version of Fullerton Advanced Balance (FAB-T) scale. *Archives of Gerontology and Geriatrics*, 78(February), 38–44. <https://doi.org/10.1016/j.archger.2018.05.022>
- Jensen, J. L., Brown, L. A., & Woollacott, M. H. (2001). Compensatory Stepping: The Biomechanics of a Preferred Response Among Older Adults. *Experimental Aging Research*, 27(4), 361–376. <https://doi.org/10.1080/03610730109342354>
- Jeon, Y.-J., & Kim, G.-M. (2017). Comparison of the Berg Balance Scale and Fullerton Advanced Balance Scale to predict falls in community-dwelling adults. *Journal of Physical Therapy Science*, 29(2), 232–234. <https://doi.org/10.1589/jpts.29.232>
- Karthikeyan, G., Sheikh, S. G., & Chippala, P. (2012). Test-retest reliability of short form of berg balance scale in elderly people. *Journal of Medicine and Medical Sciences*, 1(6), 139–144.
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–63. <https://doi.org/10.1016/j.jcm.2016.02.012>

- Kraemer, H. C., & Blasey, C. (2016). How Many Subjects?: Statistical Power Analysis in Research. EBOOK, 55 City Road, London. <https://doi.org/10.4135/9781483398761>
- Lamb, S. E., Jørstad-Stein, E. C., Hauer, K., & Becker, C. (2005). Development of a common outcome data set for fall injury prevention trials: The Prevention of Falls Network Europe consensus. *Journal of the American Geriatrics Society*, 53(9), 1618–1622. <https://doi.org/10.1111/j.1532-5415.2005.53455.x>
- Langley, F. A., & Mackintosh, S. F. H. (2007). Functional balance assessment of older community dwelling adults: a systematic review of the literature. *The Internet Journal of Allied Health Sciences and Practice*, 5(4), 1–11.
- Lee, M., Lim, T., Lee, J., Kim, K., & Yoon, B. C. (2017). Optimal retraining time for regaining functional fitness using multicomponent training after long-term detraining in older adults. *Archives of Gerontology and Geriatrics*, 73, 227–233. <https://doi.org/10.1016/j.archger.2017.07.028>
- Lee, S. H., Hee, R. N., & Kim, S. (2016). Exercise Interventions for Preventing Falls Among Older People in Care Facilities: A Meta-Analysis, 1–7. <https://doi.org/10.1111/wvn.12193>
- Mackintosh, S., Datson, N., & Fryer, C. (2006). A balance screening tool for older people: reliability and validity. *International Journal of Therapy & Rehabilitation*, 13(12), 558–561 4p. <https://doi.org/10.12968/ijtr.2006.13.12.22472>
- McIlroy, W. E., & Maki, B. E. (1996). Age-related changes in compensatory stepping in response to unpredictable perturbations. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 51(6), M289–M296. <https://doi.org/10.1093/gerona/51A.6.M289>
- Moniz-Pereira, V., Carnide, F., Ramalho, F., André, H., Machado, M., Santos-Rocha, R., & Veloso, A. P. (2013). Using a multifactorial approach to determine fall risk profiles in portuguese older adults. *Acta Reumatologica Portuguesa*, 38(4), 263–272.
- Öztürk, A., Tarsuslu, T., Yümin, E., Sertel, M., & Yümin, M. (2011). The relationship between physical , functional capacity and quality of life (QoL) among elderly people with a chronic disease. *Archives of Gerontology and Geriatrics*, 53, 278–283. <https://doi.org/10.1016/j.archger.2010.12.011>
- Park, H. A. (2013). An introduction to logistic regression: From basic concepts to interpretation with particular attention to nursing domain. *Journal of Korean*

- Academy of Nursing*, 43(2), 154–164. <https://doi.org/10.4040/jkan.2013.43.2.154>
- Pavol, M. J., Runtz, E. F., Edwards, B. J., & Pai, Y. (2002). Age Influences the Outcome of a Slipping Perturbation During Initial But Not Repeated Exposures. *Journal of Gerontology: MEDICAL SCIENCES*, 57(8), 496–503.
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1147–1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Rikli, R. E., & Jones, C. J. (1999). Development and validation of a Functional Fitness Test for Community-Residing Older Adults. *Gerontologist*. <https://doi.org/10.1093/geront/gns071>
- Rose, D. J., Lucchese, N., & Wiersma, L. D. (2006). Development of a Multidimensional Balance Scale for Use With Functionally Independent Older Adults. *Archives of Physical Medicine and Rehabilitation*, 87(11), 1478–1485. <https://doi.org/10.1016/j.apmr.2006.07.263>
- Schott, N. (2011). Erfassung der Gleichgewichtsfähigkeit bei selbstständig lebenden Erwachsenen: Reliabilität und Validität der deutschsprachigen Version der Fullerton Advanced Balance Scale. *Zeitschrift Fur Gerontologie Und Geriatrie*, 44(6), 417–428. <https://doi.org/10.1007/s00391-011-0236-8>
- Shumway-cook, A., Ciol, M. A., Hoffman, J., Dudgeon, B. J., Yorkston, K., & Chan, L. (2009). Falls in the Medicare Population : Incidence , Associated Factors , and Impact on Health Care. *Physical Therapy*, 89(4). <https://doi.org/https://doi.org/10.2522/ptj.20070107>
- Sim, J., & Wright, C. C. (2005). The Kappa Statistic in Reliability Studies : Use , Interpretation , and. *Physical Therapy*, 85(3), 257–268. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15733050>
- Todd, C., & Skelton, D. (2004). What are the main risk factors for falls amongst older people and what are the most effective interventions to prevent these falls ? *WHO Regional Office for Europe*, (March). Retrieved from <http://www.euro.who.int/document/E82552.pdf>
- Viveiro, L. A. P., Gomes, G. C. V., Bacha, J. M. R., Carvas Junior, N., Kallas, M. E., Reis, M., ... Pompeu, J. E. (2018). Reliability, Validity, and Ability to Identify Fall Status of the Berg Balance Scale, Balance Evaluation Systems Test (BESTest), Mini-BESTest, and Brief-BESTest in Older Adults Who Live in Nursing Home.

Journal of Geriatric Physical Therapy, 1.
<https://doi.org/10.1519/jpt.0000000000000215>

Wang-Hsu, E., & Smith, S. S. (2018). Interrater and Test-Retest Reliability and Minimal Detectable Change of the Balance Evaluation Systems Test (BESTest) and Subsystems with Community-Dwelling Older Adults. *Journal of Geriatric Physical Therapy*, 41(3), 173–179. <https://doi.org/10.1519/JPT.0000000000000117>

World Health Organization. (2007). WHO Global Report on Falls Prevention in Older Age. *Community Health*, 53. <https://doi.org/9789241563536>

Yuksel, E., Kalkan, S., Cekmece, S., Unver, B., & Karatosun, V. (2017). Assessing Minimal Detectable Changes and Test-Retest Reliability of the Timed Up and Go Test and the 2-Minute Walk Test in Patients With Total Knee Arthroplasty. *Journal of Arthroplasty*, 32(2), 426–430. <https://doi.org/10.1016/j.arth.2016.07.031>

Anexo I



Ficha de Registo Individual Avaliação Funcional e Equilíbrio

COD_SUJ _____

Data: _____

Composição Corporal e Força do Membro Superior

Peso (kg): _____ Altura (m): _____ IMC (kg/m²) _____ Força de preensão manual (1) _____ (Kg) (2) _____ (Kg)

Dados Antropométricos

Perímetros	Braço (cm)	Cintura (cm)	Anca (cm)	Crural (cm)	Geminal (cm)	Pregas	Tricipital (mm)	Crural (mm)	Geminal (mm)

Testes Funcionais

Levantar/sentar na cadeira: _____ (nº exec./30s) Agilidade (TUG): _____, _____ seg 2 min
Marcha: _____ (nº exec/2min)

1. Transposição do banco

- () 0 - Incapaz de colocar o apoio no banco sem perda de equilíbrio ou sem ajuda.
- () 1 - Capaz de colocar o apoio no banco com o membro inferior dominante, mas realiza interrupções em ambas as direções.
- () 2 - Capaz colocar o apoio no banco com o membro inferior dominante, mas realiza interrupções apenas numa direção.
- () 3 - Capaz de realizar a movimentação, mas requer supervisão próxima numa ou em ambas as direções.
- () 4 - Capaz de realizar a movimentação em segurança e sem ajuda em ambas as direções.

2. Dar 10 passos em linha recta

- () 0 - Incapaz de completar os 10 passos sem ajuda.
- () 1 - Capaz de completar os 10 passos com seis ou mais interrupções.
- () 2 - Capaz de completar os 10 passos com três a cinco interrupções.
- () 3 - Capaz de completar os 10 passos com uma ou duas interrupções.
- () 4 - Capaz de completar os 10 passos sem ajuda e sem interrupções.

3. Equilíbrio unipedal

- () 0 - Incapaz de tentar, ou necessita de ajuda para prevenir a queda.
- () 1 - Capaz de elevar o membro inferior sem ajuda, mas incapaz de manter a posição mais de 5 segundos (≤ 5 s).
- () 2 - Capaz de elevar o membro inferior sem ajuda, e mantém a posição mais de 5 mas ≤ 12 segundos.
- () 3 - Capaz de elevar o membro inferior sem ajuda e de manter a posição mais de 12 mas < 20 segundos.

- () 4 - Capaz de elevar o membro inferior sem ajuda e de manter a posição durante 20 segundos.

4. *Equilíbrio na superfície de espuma*

- () 0 - Incapaz de realizar e manter a posição na superfície de espuma sem ajuda e de manter os olhos fechados.
- () 1 - Capaz de realizar e manter a posição na superfície de espuma, mas incapaz ou pouco disposto a fechar olhos.
- () 2 - Capaz de realizar e manter a posição na superfície de espuma, com os olhos fechados durante ≤ 10 segundos.
- () 3 - Capaz de realizar e manter a posição na superfície de espuma com os olhos fechados > 10 e < 20 segundos.
- () 4 - Capaz de realizar e manter a posição na superfície de espuma com os olhos fechados durante 20 segundos.

FCT Fundação para a Ciência e a Tecnologia

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Projeto PTDC/DES/72946/2006

Projeto EXPL/DTP/DES/1915/2013